

# Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/GB05/000937

International filing date: 10 March 2005 (10.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: GB  
Number: 0407816.8  
Filing date: 06 April 2004 (06.04.2004)

Date of receipt at the International Bureau: 13 April 2005 (13.04.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland  
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse



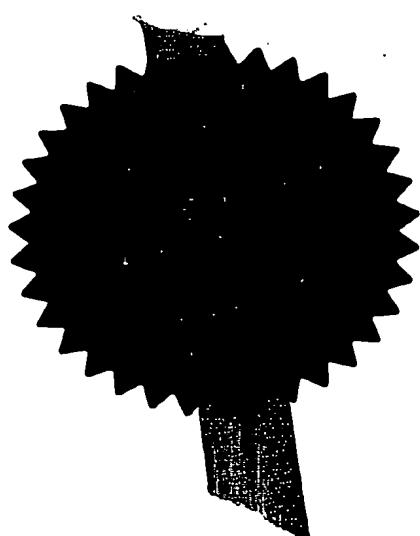
The Patent Office  
Concept House  
Cardiff Road  
Newport  
South Wales  
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.



Signed

Dated 31 March 2005



9. Enter the number of sheets for any of the following items you are filing with this form.  
Do not count copies of the same document

Continuation sheets of this form

Description	13
Claim(s)	2
Abstract	1
Drawing(s)	2 + 2 figs

10. If you are also filing any of the following, state how many against each item

## Priority Documents

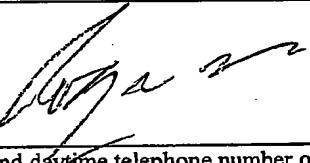
Translations of priority documents

Statement of inventorship and right to grant of a patent (Patents Form 7/77)

Request for preliminary examination and search (Patents Form 9/77)

Request for substantive examination (Patents Form 10/77)

Any other documents (please specify)

11.  I/We request the grant of a patent on the basis of this application.  
Signature(s) Date:  
**06 April 2004**  
**NASH, Roger William Authorised Signatory**

12. Name and daytime telephone number of person to contact in the United Kingdom **Mark Watson** **020 7356 6163**

## Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

## Notes

- If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.
- Write your answers in capital letters using black ink or you may type them.
- If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- If you have answered 'Yes' Patents Form 7/77 will need to be filed.
- Once you have filled in the form you must remember to sign and date it.
- For details of the fee and ways to pay please contact the Patent Office.

## INFORMATION RETRIEVAL

The present invention relates to an information retrieval apparatus and method.

According to a first aspect of the present invention there is provided a method for accessing an information resource, comprising the steps of:

- (i) receiving a user query;
- (ii) comparing portions of the user query with phrases in a set of predefined phrases to find one or more matching phrases;
- (iii) identifying, using predefined relationships between said predefined phrases and predefined concepts in an ontology, one or more concepts relevant to said portions of the received user query; and
- (iv) identifying, using predefined relationships between predefined actions and said predefined concepts, one or more actions relevant to the received user query, wherein an action comprises providing access to an information resource.

Preferably, said predefined concepts comprise task concepts and non-task concepts, and the ontology defines, for each task concept, an indication of the number of non-task concepts required to implement a corresponding task.

In a preferred embodiment of the present invention, there is provided a further step:

- (vi) in the event that said one or more concepts identified at step (iii) are insufficiently specific to enable a relevant action to be identified at step (iv), identifying from the ontology one or more further concepts related to those identified at step (iii) and requesting input from a user to select one or more of said further concepts for use in step (iv) to identify a relevant action.

Apparatus according to the present invention may be applied as a "just-in-time" information assistant which uses an ontology to improve the management and selection of information to be displayed to a user. In addition to supplying information, preferred embodiments of the present invention enable user queries to be linked to business processes and people. For example, in a contact centre application the apparatus accepts an incoming message, e.g. an operator dialogue with a customer or an email, and matches the message to concepts in the ontology. Combinations of these matched concepts are then used to show information, select a business process or locate a relevant person.

The ontology is a representation of relevant entities along with important properties and their relationships. For example the products supplied by a company are the relevant entities whilst information about which are EEC compliant are important properties. In preferred embodiments of the present invention the ontology is implemented as a hierarchy in which child nodes are instances of a parent node. The ontology enables reuse of defined concepts for different domains of application and enables task-related concepts, e.g. fault, pricing information, to be identified separately from entities such as product types.

It is not just documents which can be attached to entities in the ontology, but also processes and people. A call centre operator for example may therefore be directed more quickly to the correct response in respect of a customer enquiry, i.e. relaying a piece of information, activating the correct business process or contacting the correct person.

Two interactive modes or operation of the apparatus are supported according to preferred embodiments of the present invention: in one mode the apparatus is able to carry on a dialogue with a user in order to resolve a query that is too broad; in another mode the apparatus may monitor telephonic or instant messaging conversations between a customer and a call centre operator, for example, analysing the conversation to continuously identify key concepts in the conversation and to construct relevant queries to automatically supply information, identify processes or people relevant to the subject matter being discussed with the customer.

Preferred embodiments of the present invention use an ontology:

- (1) To organise resources such as documents, business processes and domain experts. It effectively provides a concept-based indexing to these resources. As the ontology is formal and highly structured, it allows fast and accurate resource retrieval using structured queries instead of merely generating a list of hits as is often returned by known answer engines.
- (2) To help analyse the correct intention of a user query. The invention's dialogue module uses relationships and constraints for each of the defined concepts to ascertain relevant tasks which may apply.

Fuzzy techniques are used to map concepts in the ontology to words and phrases likely to arise in user queries and hence to handle the idiosyncrasies and unstructured nature of user queries.

According to a second aspect of the present invention there is provided an information retrieval apparatus, comprising:

an input for receiving a user query;

an ontological database for storing an ontology defining relationships between a plurality of predefined concepts;

a context phrase database for storing predefined context phrases and, for each context phrase, information defining a fuzzy relationship with an associated concept stored in the ontology;

a concept mapper for comparing portions of a received user query with context phrases stored in the context phrase database to thereby identify and output one or more relevant concepts; and

an action selector operable to identify an action in respect of one or more relevant concepts output by the concept mapper, wherein an action comprises providing access to an information resource in response to the received user query.

Preferred embodiments of the present invention will now be described in more detail, by way of example only, with reference to the accompanying drawings of which:

Figure 1 is a diagram showing features of an apparatus according to preferred embodiments of the present invention; and

Figure 2 is a flow diagram showing steps in operation of a fuzzy concept mapper according to a preferred embodiment of the present invention.

A preferred apparatus and its operation according to a preferred embodiment of the present invention will now be described in overview with reference to Figure 1.

Referring to Figure 1, the apparatus 100 is provided with a query input 105 arranged to receive a query from a user. Of course, a user query need not be an actual question. In a preferred call centre application of the present invention, it may be appropriate simply to ensure that relevant information is always available on-screen to the call centre operator (user of the apparatus 100) while processing a customer enquiry. On receipt of a new query at the query input 105 a new query session is initiated within the apparatus 100. The query input 105 is arranged to receive a user query by a number of different channels. For example, the query may be received in the form of an e-mail message or as a natural language query submitted by means of a web page or an instant messaging interface. Alternatively, speech recognition software may be used to convert a user's spoken dialogue into a text input to the query input 105, in real time, for processing by the apparatus 100 as the dialogue progresses.

Once a query text has been received at the query input 105, or while text is being received, it is passed to a so-called "phrase chunker" 110. The phrase chunker 110 separates input queries into smaller chunks, i.e. phrases which can be matched to

concepts. Preferably, the phrase chunker 110 is arranged to divide the received query text into n-grams - sequences of n words or fewer, ideally with  $n < 5$  – wherein an n-gram does not cross a sentence boundary. Alternatively, the phrase chunker may operate according to a known yet more sophisticated algorithm, designed to identify phrases of up to a predetermined length comprising words more likely to be indicative of the concepts embodied in the user query, eliminating certain “low value” words before constructing those phrases for example.

Output from the phrase chunker 110 is submitted to a fuzzy concept mapper 115 operable to identify one or more predefined concepts stored in an ontology database 120 that appear to have the greatest relevance to terms and phrases output from the phrase chunker 110. The fuzzy concept mapper 115 identifies concepts by firstly looking for context phrases stored in a context phrase database 125 that match terms and phrases contained in the query input. Predefined fuzzy relationships are maintained between concepts stored in the ontology database 120 and context phrases stored in the context phrase database 125. Therefore, having identified one or more matching context phrases (125), the fuzzy concept mapper 115 is able to identify one or more relevant concepts by analysing the respective fuzzy relationships. A more detailed description of the operation of the fuzzy concept mapper 115 will be provided below.

The fuzzy concept mapper 115 is arranged to generate and to update a list of the current concepts identified in a received user query at any one time. For example, if the user query is being captured from dialogue, the fuzzy concept mapper 115 is arranged to continually look for relevant concepts as query text is received (105) and processed by the apparatus 100, to add newly identified concepts to the current concept list and to update fuzzy support values (relevance weightings) associated with those concepts already identified. It is therefore important that when a new user query is received at the query input 105, or when it is otherwise determined that the apparatus 100 should be reset with respect to an ongoing user query, that the list of current concepts is emptied.

The fuzzy concept mapper 115 looks in the ontology (120) for relevant concepts of two types: task and non-task. The ontology (120) defines for each task concept the number and type of non-task concepts that would be required to fully define the task. The fuzzy concept mapper 115 is therefore arranged to recognise an event in which a task concept and a required number of non-task concepts has been identified in respect of a given user query and, at this point, to output the current

given concept typically comprises a parent concept and a set of one or more child

concepts. Preferably, the ontology distinguishes task concepts from non-task concepts. Task concepts are abstract tasks, e.g. fault, sales, pricing, overview, etc. Each concept may have associated with it a set of one or more properties. In particular, a non-task concept may have a property that defines, for example, whether specific task concepts can be associated with it.

By way of example, a section of an ontology as may be stored in the ontology database 120 comprises a hierarchy of concepts, as follows,:

#### TASKS

- Describe\_Benefits
- Pricing
- Buy
- Fault
- Reconnect
- Information
- Alter\_details
- Compare
  - prices
  - features

#### PRODUCTS

- PHYSICAL-PRODUCTS
  - CORDLESS-PHONES
  - ANSWERING-MACHINES
  - FAXES
- INTERNET-ACCESS
  - DIAL-UP
  - MIDBAND
  - BROADBAND
- PSTN
  - Friends&Family

In this example, there are two types of concept in the ontology: "TASKS" and "PRODUCTS." The ontology is arranged in a hierarchical fashion with TASKS and PRODUCTS being the root nodes of the ontology. Each "child" node under the "parent"

information. A known algorithm may be used to extract key terms from an input

document and/or to suggest where in the hierarchy of the ontology (120) a concept should be placed and which context phrases should be associated with it.

For each concept defined in the ontology database 120 there is provided, in the context phrase database 125, an associated list of key phrases which are related to the concept. A fuzzy measure of support between 0 and 1 is recorded against each key phrase, indicative of the relevance of the phrase to the associated concept. For example, for the concept *task:fault*; the relevant key phrases and measures of support that might be recorded in the context phrase database 125 are:

broken: 0.9

not working: 0.9

loose: 0.3

squeeky: 0.1

The context phrases selected for inclusion in the context phrase database 125 are those phrases most likely to be used in user queries. The context phrase database 125 therefore provides a link between terms that might be expected to occur in a typical user query and concepts defined in the ontology (120). This link is exploited by the fuzzy concept mapper 115 in order to identify, by comparing portions of a received user query that have been output by the phrase chunker 110 with stored context phrases (125), one or more concepts of greatest relevance to the received user query. Preferred steps in operation of the fuzzy concept mapper 115 for identifying one or more concepts of relevance to a new user query will now be described with reference to Figure 2. The process to be described may operate to analyse a user query that has been received complete, e.g. in the form of an e-mail, or to analyse portions of a user query as it is being received, e.g. during an ongoing conversation between a call centre operator and a customer.

Referring to Figure 2, the preferred process begins at STEP 200 by initialising the current concept list for the user query so that the process begins with an empty list, or a list comprising one or more default concepts with associated fuzzy support values. A portion of the user query is received at STEP 205 from the phrase chunker 110. At STEP 210 the received portion is compared with context phrases stored in the context phrase database 125. If, at STEP 215, no matching context phrases are found, then processing proceeds to STEP 250 to determine whether the end of the user query

list for the current task. If there are more than n other concepts in the list, the action

---

the resultant web page to the user. All content may be highly structured and represented

choice must be made between broadband and mid-band in order to define what is

---

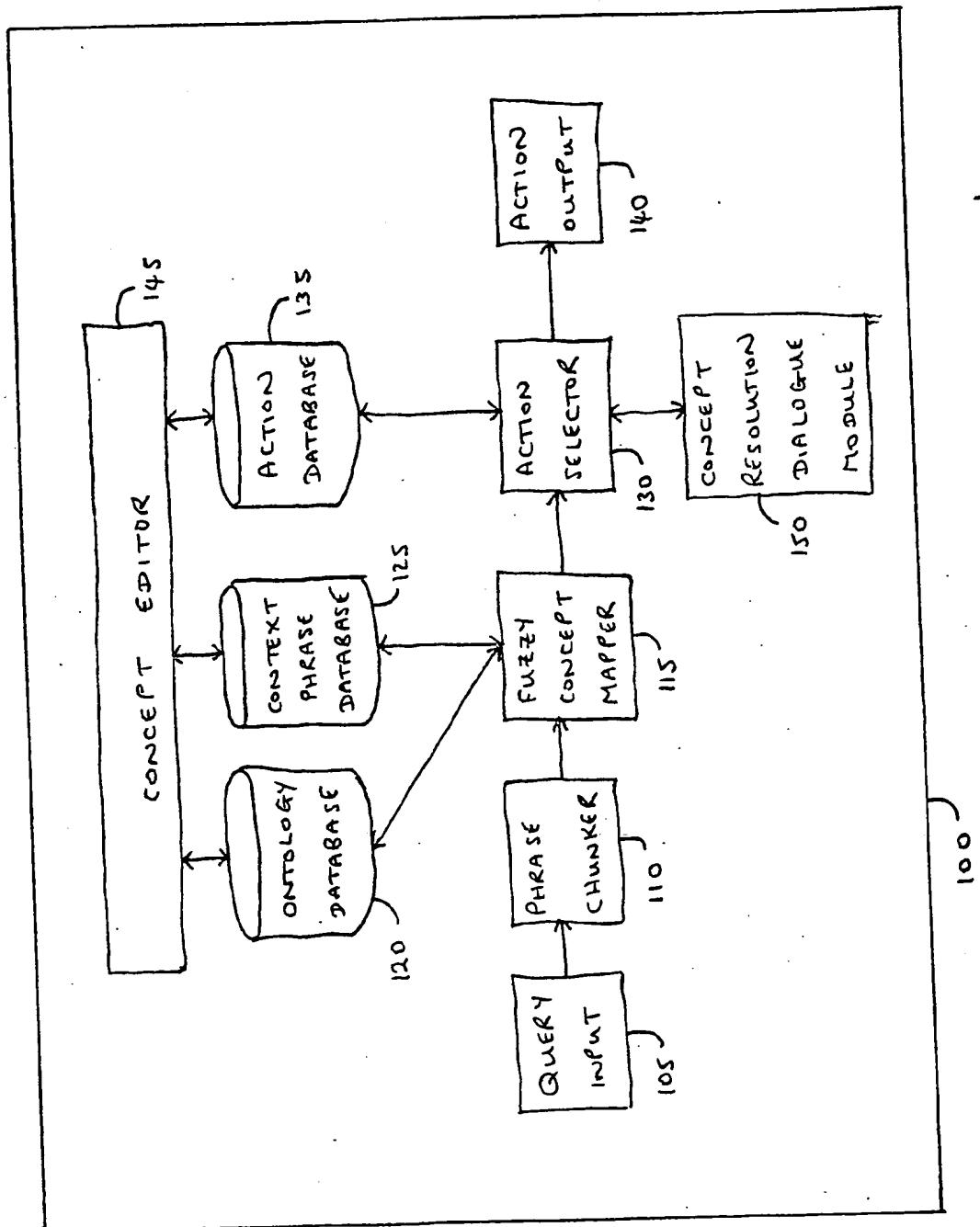


FIGURE 1

2/2

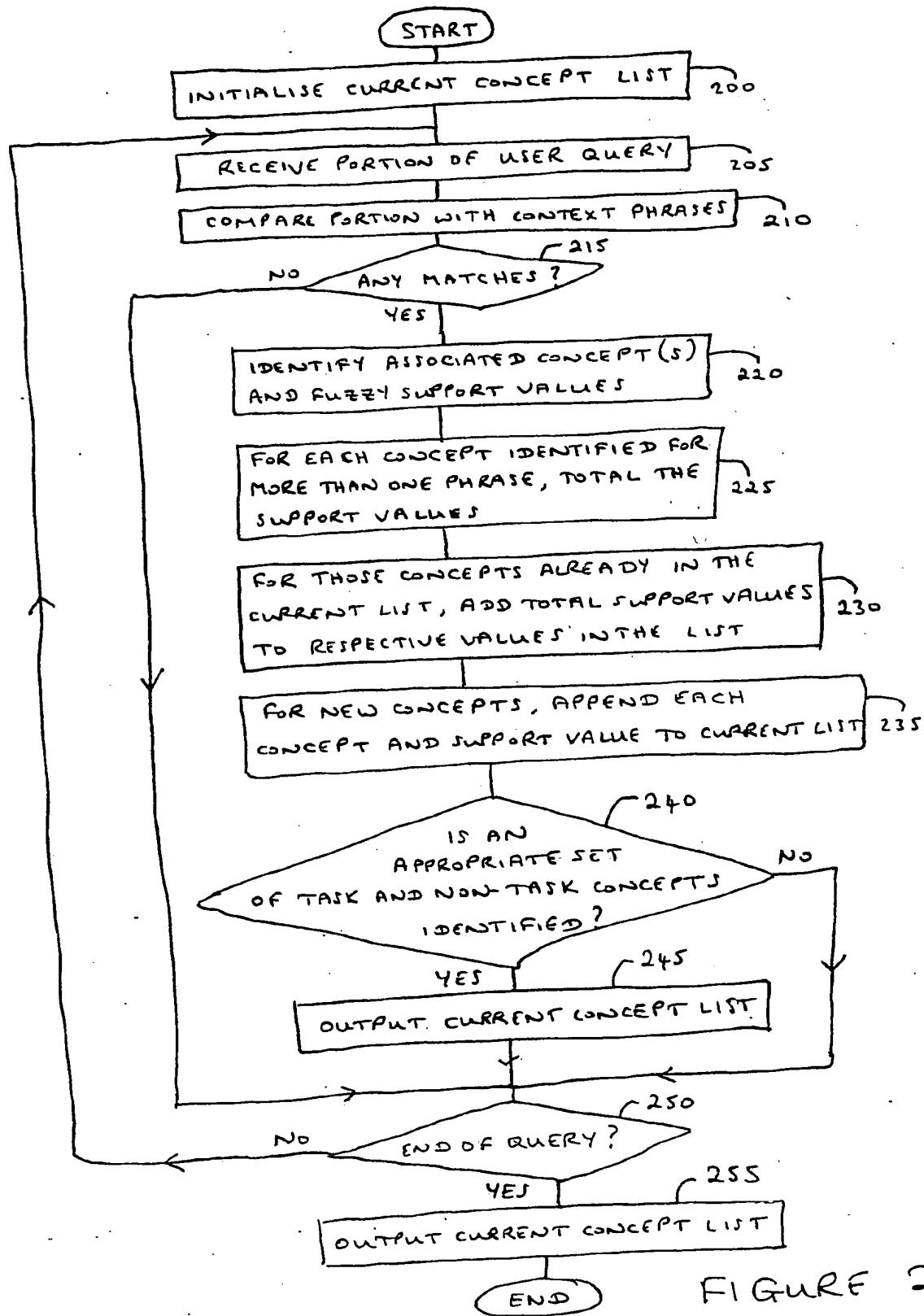


FIGURE 2

From the INTERNATIONAL BUREAU

**PCT**

NOTIFICATION CONCERNING  
SUBMISSION OR TRANSMITTAL  
OF PRIORITY DOCUMENT

(PCT Administrative Instructions, Section 411)

To:

LAU, Chi-Fai  
BT Group Legal, Intellectual Property Department  
PP: C5A, BT Centre  
81 Newgate Street  
Greater London  
LONDON EC1A 7AJ  
ROYAUME-UNI

Date of mailing (day/month/year) 21 April 2005 (21.04.2005)			
Applicant's or agent's file reference A30476 WO	<b>IMPORTANT NOTIFICATION</b>		
International application No. PCT/GB05/000937	International filing date (day/month/year) 10 March 2005 (10.03.2005)		
International publication date (day/month/year)	Priority date (day/month/year) 06 April 2004 (06.04.2004)		
Applicant BRITISH TELECOMMUNICATIONS PUBLIC LIMITED COMPANY et al			

1. By means of this Form, which replaces any previously issued notification concerning submission or transmittal of priority documents, the applicant is hereby notified of the date of receipt by the International Bureau of the priority document(s) relating to all earlier application(s) whose priority is claimed. Unless otherwise indicated by the letters "NR", in the right-hand column or by an asterisk appearing next to a date of receipt, the priority document concerned was submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b).
2. *(If applicable)* The letters "NR" appearing in the right-hand column denote a priority document which, on the date of mailing of this Form, had not yet been received by the International Bureau under Rule 17.1(a) or (b). Where, under Rule 17.1(a), the priority document must be submitted by the applicant to the receiving Office or the International Bureau, but the applicant fails to submit the priority document within the applicable time limit under that Rule, the attention of the applicant is directed to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.
3. *(If applicable)* An asterisk (\*) appearing next to a date of receipt, in the right-hand column, denotes a priority document submitted or transmitted to the International Bureau but not in compliance with Rule 17.1(a) or (b) (the priority document was received after the time limit prescribed in Rule 17.1(a) or the request to prepare and transmit the priority document was submitted to the receiving Office after the applicable time limit under Rule 17.1(b)). Even though the priority document was not furnished in compliance with Rule 17.1(a) or (b), the International Bureau will nevertheless transmit a copy of the document to the designated Offices, for their consideration. In case such a copy is not accepted by the designated Office as the priority document, Rule 17.1(c) provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.

Priority date	Priority application No.	Country or regional Office or PCT receiving Office	Date of receipt of priority document
06 April 2004 (06.04.2004)	0407816.8	GB	13 April 2005 (13.04.2005)

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland  Facsimile No. +41 22 740 14 35	Authorized officer  Boccard Paulette  Facsimile No. +41 22 338 87 40 Telephone No. +41 22 338 8147
---	---